Impact of paternity errors in cow identification on genetic evaluations and international comparisons

J. Dairy Sci. 84:2523–2529

Impact of Paternity Errors in Cow Identification on Genetic Evaluations and International Comparisons

G. Banos,*[,1] G. R. Wiggans,† and R. L. Powell†
*Interbull Centre, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden
†Animal Improvement Programs Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

ABSTRACT

The impact of paternity identification errors on US genetic evaluations and international comparisons of Holstein dairy bulls for milk, fat, and protein yields was investigated. Sire identification was replaced for 11% of Holstein cows that were sired by AI bulls and had records in the US database for national genetic evaluations; US evaluations were computed based on those modified pedigrees and compared with official national evaluations. Estimated breeding values from the data with introduced paternity errors were biased, especially for later generations. Estimated genetic trends decreased by 11 to 15%. Estimates of standard deviations of sire transmitting ability also decreased by 8 to 9%. International multitrait across-country comparisons of bulls were computed based on national evaluations from the United States, Canada, New Zealand, and The Netherlands. Estimates of genetic correlations between the United States and other countries decreased by 0.04 to 0.06 when US evaluations were based on modified pedigree. The resulting bias toward selection of domestic bulls and the inability to identify truly superior animals that are available internationally could decrease potential selection differentials by 0.07 to 0.09 standard deviation units on the US scale, which corresponds to sire breeding values of approximately 50 kg for milk, 3 kg for fat, and 1.7 kg for protein. Losses for the other countries were lower and ranged from 0.02 to 0.05 standard deviation units, because a correlation of less than unity with the United States decreased the impact of US cow paternity errors on the scales of other countries. Although paternity verification is desirable and technically feasible, commercial implementation would require low testing costs.

(**Key words:** pedigree error, genetic evaluation, international comparison, dairy cattle)

INTRODUCTION

Genetic evaluations are computed to assess the genetic merit of dairy cattle. Elaborate animal models have been developed for the analysis of performance records and genetic relationships among individual animals. An important assumption for those models is that all identified genetic relationships are correct. However, the literature suggests that this assumption is incorrect. Reported rates of paternity misidentification in various cattle populations worldwide reached almost 25%—around 5% in Israel (Ron et al., 1996), around 12% in The Netherlands (Bovenhuis and van Arendonk, 1991), 5 to 15% in Denmark (Christensen et al., 1982), 8 to 20% in Ireland (Beechinor and Kelly, 1987), and 4 to 23% in Germany (Gelderman et al., 1986).

Paternity errors are expected to bias estimation of genetic parameters (Van Vleck, 1970b), breeding values of individual animals (Israel and Weller, 2000; Van Vleck, 1970a), and genetic progress (Gelderman et al., 1986; Van Vleck, 1970a). Van Vleck (1970b) also showed that bias increased as the proportion of records with errors increased. Those studies were based on either deterministic analysis that considers the sire-daughter inheritance path (Gelderman et al., 1986; Van Vleck, 1970a and 1970b) or stochastic simulation of commercial cattle populations and animal model evaluations (Israel and Weller, 2000). In the latter study, an error rate of 10% was simulated for cow paternity and resulted in decreased genetic gain and biased EBV.

Previous studies investigated the impact of cow paternity errors on genetic evaluations of cattle in the context of a national breeding program. Currently, national genetic evaluation results are also used as input to international genetic evaluations and comparisons of dairy bulls (Banos and Sigurdsson, 1996; Schaeffer, 1994). Biased national genetic evaluations affect such international comparisons. The impact of cow paternity

Received November 6, 2000.
Accepted June 6, 2001.
Corresponding author: Georgios Banos; e-mail: banos@vet.auth.gr.
[1]Current address: Faculty of Veterinary Medicine, Aristotle University, Box 393, GR-54006 Thessaloniki, Greece.

errors in one country on international genetic comparisons has yet to be investigated.

The objective of this study was to determine the effect of misidentification of cow sires on genetic evaluations and estimation of genetic parameters and breeding values for bulls across countries.

MATERIALS AND METHODS

Sire identification was replaced for a random sample of 11% of Holstein cows that were sired by AI bulls and had records in the US database for national genetic evaluations. An error rate of 11% was the average paternity error rate reported in the international scientific literature. The replacement sire was another bull of similar age that had been born in the same, previous, or following year as the true sire. Replacement was also within category of AI service status: 1) bulls that had returned to service as indicated by >150 daughters and >5 yr in age difference from the cow with the paternity change, or 2) bulls that were being progeny tested as well as a few other AI bulls that were not returned to service. Assignment of false daughters was random among eligible bulls without regard to the number of true daughters. A total of 1,574,244 cows had their sire identification replaced.

Genetic evaluations for milk, fat, and protein yields were computed with the animal model by using USDA evaluation procedures, the modified pedigree data, and cow lactation information that had been included in the official US evaluations of May 2000. Resulting evaluations were compared with official May 2000 results to assess the impact of the introduced paternity errors on national EBV.

The international comparison included official May 2000 national evaluations from Canada, New Zealand, and The Netherlands in addition to US bulls that had both official May 2000 US evaluations and May 2000 US evaluations based on modified cow pedigree. Data edits followed the standards of the International Bull Evaluation Service (2000), and AI Holstein bulls born during 1983 or later with daughters in ≥10 herds in the country of origin were included. International genetic evaluations for milk, fat, and protein yields were computed using a multitrait across-country evaluation procedure (Schaeffer, 1994) and official May 2000 national evaluations from Canada, New Zealand, The Netherlands, and the United States. A second set of international evaluations was computed with official US evaluations replaced by US evaluations that were based on modified pedigrees. Results from the two international evaluations were compared to determine the impact that cow paternity errors in one country have on international genetic comparisons.

RESULTS

Mean changes in numbers of daughters per US Holstein sire after the introduction of cow paternity errors at 11% are shown in Table 1. As expected, mean losses in true daughters were compensated by mean gains in false daughters. For individual bulls, however, losses and gains were not always equal. Standard deviations for losses were higher than for gains for all service categories of bulls because the number of lost daughters was proportional to the true number of daughters of a bull, whereas each bull always had the same chance to have new daughters assigned regardless of the number of true daughters.

National Evaluations

An 11% introduction of cow paternity errors changed some existing genetic relationships among animals that had resulted from family selection and affected estimated inbreeding. Mean estimated inbreeding of all animals was reduced by 7 to 11% (Table 2). The impact on younger animals was more severe because more historic genetic relationships contribute to later generations. These estimates, however, are not correct predictors of true inbreeding because an animal's inbreeding is determined by the genetic relationship of its true parents. To the extent that inbreeding is related to selection and the latter is affected by pedigree accuracy (see next paragraph), true inbreeding would be also affected by sire misidentification.

Table 1. Changes in daughter counts from random replacement of cow paternity identification.

Category of AI service status	Bulls, n	True daughters, no.		True daughters lost, no.		False daughters gained, no.	
		Mean	SD	Mean	SD	Mean	SD
All bulls	32,749	432	2083	47.8	230.8	47.8	95.7
Returned to service[1]	5655	2230	4608	247.9	511.3	247.9	84.5
Progeny test[2]	27,094	57	27	6.7	4.2	6.7	2.8

[1]Bulls with >150 daughters and >5 yr difference in age from daughters.

[2]All other bulls that were not considered to have returned to service.

Table 2. Mean estimates of inbreeding in the US Holstein population with introduced cow paternity errors.

Birth year	Gender	Mean inbreeding		Reduction in inbreeding from increase in paternity errors
		0% errors	11% errors	
				(%)
All years	Cows	0.0169	0.0158	6.5
	AI bulls	0.0175	0.0156	10.9
1990 or later	Cows	0.0352	0.0321	8.8
	AI bulls	0.0421	0.0363	13.8

The incorrect pedigree information affected EBV, especially for bulls of high and low genetic merit. New reputed daughters assigned to high-merit bulls through random replacement were expected to be of generally lower merit than the true progeny that were replaced. Estimated annual genetic trends decreased by 11 to 15% overall (Table 3), which was similar to results from the simulation study of Gelderman et al. (1986) but more than found through simulation by Israel and Weller (2000). The observed decrease in annual trend was larger than Israel and Weller's (2000) probably because truly top bulls would not have been used as heavily here, had selection been based on their biased genetic evaluation. Bias was more evident for later generations for which estimates of annual genetic trend decreased the most. Estimated trend for all bulls decreased by 14 to 15%, whereas the decrease was up to 18% for bulls born during 1983 or later.

Since sire replacement was by a bull of similar age, a possible explanation for the compounding impact of misidentification on genetic trend over time is that more recent sires have a greater proportion of misidentified relatives than do base parents. A typical base bull would have 11% of daughters but not any older relatives that were misidentified. Conversely, a young bull might have many relatives in addition to daughters that had been misidentified so that the total proportion of misidentified relatives increased over generations.

True genetic progress in the US population was not affected, because actual selection had been based on official evaluations that had been estimated without introduced pedigree errors. Thus, the reduced trends from this study reflected the impact of erroneous information on evaluation variance and regression slope. Reduction in variability of genetic evaluations was 7 to 11% after introduction of random cow paternity errors. Figure 1 illustrates this observation with means and SD of EBV for protein yield for AI Holstein bulls born during 1983 or later with daughters in ≥10 herds; those bulls subsequently were evaluated internationally.

International Comparisons

Numbers of Holstein bull records from each country and estimates of within-country SD for sire transmitting abilities for yield traits are shown in Table 4 for the four-country international comparison. Estimates of sire SD decreased by 8 to 9% for US transmitting abilities based on data with an 11% error rate for cow paternity, which reflected the reduced variability of the

Table 3. Genetic trend in EBV for yield traits of the US Holstein population with introduced cow paternity errors.

Gender	Yield trait	Trend in EBV		Reduction in trend from increase in paternity errors
		0% errors	11% errors	
		(kg/yr)		(%)
Cows	Milk	116.3	103.6	11.1
	Fat	4.08	3.63	11.0
	Protein	3.69	3.22	12.7
AI bulls				
All bulls	Milk	123.2	106.0	13.9
	Fat	4.11	3.52	14.3
	Protein	4.19	3.55	15.2
US bulls with international evaluations[1]	Milk	113.8	93.3	18.0
	Fat	3.97	3.39	14.4
	Protein	4.35	3.65	16.0

[1]Born during 1983 or later with daughters in ≥10 herds.

Table 4. Numbers of bull records and estimates of sire SD for EBV of yield traits in the international comparison.

Country	Bull records, n	Sire SD for EBV		
		Milk	Fat	Protein
		(kg)		
United States				
Official evaluations[1]	16,526	339	12.3	9.4
Evaluations with introduced pedigree errors[2]	16,526	309	11.2	8.6
Canada[1]	4697	396	15.2	11.3
New Zealand[1]	2452	157	6.2	4.6
The Netherlands[1]	5182	271	10.4	7.4

[1]Data from official May 2000 national evaluations.

[2]Data from official May 2000 national evaluations with an 11% introduced error rate for cow paternity.

national evaluations (Figure 1). Sire SD is an important parameter in international genetic evaluations because it affects the slope of EBV conversion from one country to another (International Bull Evaluation Service, 2000). Decreases in sire SD of 8 to 9% would affect international rankings of dairy bulls.

The introduction of random paternity errors at a rate of 11% also resulted in decreases of 0.04 to 0.06 in estimated genetic correlations between the United States and the other countries (Table 5). Genetic correlations were estimated with the method of Sigurdsson et al. (1996), which depends on the presence of genetic links between countries. Although the numbers of common bulls and other links between countries were substantial and remained the same regardless of introduced cow paternity errors, the reduced variability and trend and other changes in the US evaluations that were caused by the introduced errors affected the estimates of genetic correlations. Unfortunately, no appropriate procedure to calculate standard errors of the estimates was available. Because of the high level of data connectedness and links among the populations, the estimates of genetic correlations were considered to be accurate. Furthermore, correlation estimates of the United States with all other countries decreased by about the same amount. Decreased estimates of genetic correlations would affect international EBV.

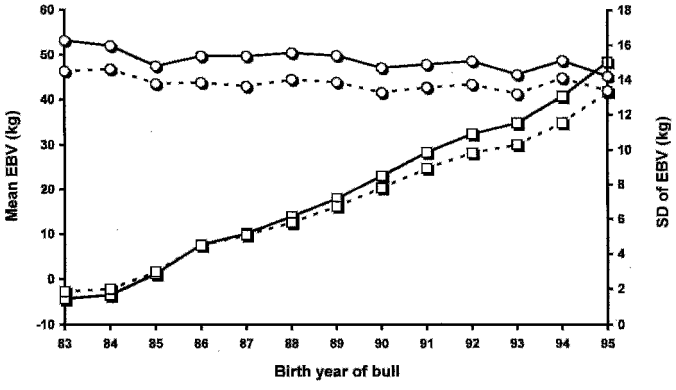

Figure 1. Mean (□) and SD (○) of protein EBV of US AI bulls from national genetic evaluations with introduced cow paternity errors at a rate of 0% (——) and 11% (– – –) by birth year of bull.

Within year product-moment correlations were calculated between international genetic evaluations that included official US evaluations and those that included US evaluations with introduced paternity errors (Table 6). Rank correlations also were computed and were found to be 0 to 2% lower than product-moment correlations. Across all bulls in the international comparison, correlations were about 0.98, which indicated that some minor reranking of bulls could be expected from introduced US pedigree errors. Results for US bulls were affected the most; correlations among international EBV decreased to about 0.95. Consequently, comparisons between US bulls and those from other countries would be expected to change as the result of US paternity errors. Correlations among international EBV for bulls from the other countries were close to unity, which indicated that comparisons among those bulls were unchanged by introduced US pedigree errors.

Changes in estimates of variances of sire EBV and genetic correlations affect international conversions and bull comparisons between countries. In general, a low estimate of the genetic correlation between countries favors local bulls regardless of scale. Therefore, the introduction of cow paternity errors in US pedigrees results in more domestic bulls among the top bulls for genetic merit on the evaluation scale of each participating country.

Bias toward the selection of domestic bulls and an inability to identify animals that are truly genetically superior and available internationally decreases potential genetic gains. Selection differentials for international evaluations that included official US evaluations or US evaluations with introduced pedigree errors were

Table 5. Estimated genetic correlations between the United States and other countries for official May 2000 US evaluations for yield traits and US evaluations with an introduced error rate of 11% for cow paternity.

	Milk yield		Fat yield		Protein yield	
Country	Official US evaluations	US evaluations with pedigree errors	Official US evaluations	US evaluations with pedigree errors	Official US evaluations	US evaluations with pedigree errors
Canada	0.95	0.89	0.95	0.90	0.94	0.88
New Zealand	0.78	0.73	0.78	0.74	0.76	0.72
The Netherlands	0.93	0.88	0.92	0.88	0.94	0.89

estimated based on genetic merit of the best 75 bulls. Selection of the top 75 bulls is consistent with the total number of bull sires used annually in the four countries (considering sires with >3 sons). With an introduced error rate of 11% for cow paternity, selection differentials decreased by 0.07 to 0.09 genetic SD units on the US scale (Table 7), which corresponds to a loss of approximately 50 kg of milk, 3 kg of fat, and 1.7 kg of protein for sire breeding value. Selection differential losses for the other countries were lower and ranged from 0.02 to 0.05 genetic SD units, because a genetic correlation of less than unity with the United States decreased the impact of US cow paternity errors on the scales of those countries. New Zealand had the lowest genetic correlation with the United States and had the lowest selection differential losses.

DISCUSSION

Sires of cows may be misidentified for many reasons. The 11% rate of introduced pedigree errors in this study was based on the incidence of errors in the current global population. Sire identifications were reassigned randomly within groups of bulls of similar age and AI service status. Consequently, daughters of sires that had returned to service would be distributed among other proven, highly used bulls. Among young bulls, however, high-merit sires could have been replacing bulls of either high or low merit. An alternative, more complicated paternity replacement strategy would have been to reassign sire identification according to the genetic merit of the true sire of a cow so that daughters of young bulls of high merit would have been distributed among other sires of high merit. The full impact of such a strategy on genetic evaluations is not intuitively clear. Because young sires of exceptionally high merit still would be replaced by a sire of lower merit, genetic variance and trend still would be expected to decrease, thereby affecting national and international genetic evaluations.

This study was based on a single reanalysis of the US national genetic evaluation. Replication of several sets of introducing random parentage error would have allowed estimation of the standard error of the observed bias. This, although desirable, was not attainable from a practical point of view, as it would have required several reruns of the entire US genetic evaluation system. However, the study yielded some useful results whose consistency with results from previous simulation studies may attest to the validity of the design.

In this study, all daughters of known AI Holstein sires had an equal chance of having their parentage

Table 6. Product-moment correlations between May 2000 international genetic evaluations that include official US evaluations for yield traits and US evaluations with an introduced error rate of 11% for cow paternity.

		Yield trait		
Country of bull	Country scale	Milk	Fat	Protein
All bulls	United States	0.975	0.971	0.970
	Canada	0.981	0.978	0.977
	New Zealand	0.983	0.978	0.980
	The Netherlands	0.981	0.978	0.976
United States	United States	0.953	0.957	0.949
	Canada	0.955	0.959	0.950
	New Zealand	0.959	0.960	0.957
	The Netherlands	0.955	0.959	0.950
Other than the United States	United States	0.996	0.996	0.994
	Other[1]	0.999	0.999	0.999

[1]Mean of correlations on the scales of Canada, New Zealand, and The Netherlands.

Table 7. Selection differential losses from use of the top 75 bulls for yield traits based on May 2000 international evaluations that include US evaluations with an introduced error rate of 11% for cow paternity rather than official US evaluations.

	Yield trait		
Country scale	Milk	Fat	Protein
	(SD units)		
United States	0.071	0.078	0.092
Canada	0.047	0.026	0.036
New Zealand	0.028	0.016	0.022
The Netherlands	0.037	0.038	0.053

changed; those daughters included influential bull dams with many sons as well as embryo transfer donor dams of high merit. Donor dams, however, may actually have incorrect pedigrees less often than do progeny-test daughters. If this situation were true, results from this study would be more pertinent to progeny-test bulls than to older sires of bull dams.

The same heritability was assumed for both national and international US evaluations, even after cow pedigree errors had been introduced. Incorrect pedigree information may bias the estimation of heritability downward depending on the fraction of cows with incorrect paternity (Van Vleck, 1970b). A heritability estimate that was closer to 0.24 instead of 0.30, the current heritability that is assumed in the US evaluation system (International Bull Evaluation Service, 2000), might have been more appropriate. However, Christensen et al. (1982) concluded from their simulation study that the theoretical heritability estimate is still appropriate even in the presence of pedigree errors. Consequently, in the simulation study of Israel and Weller (2000), the same heritability estimate was used in both the presence and the absence of pedigree errors.

The introduction of cow paternity errors at the rate of 11% decreased estimates of genetic correlations among countries by 0.04 to 0.06. For international evaluations, genetic correlations are interpreted as the interaction of genotype with environment; lower correlation estimates indicate higher interaction effects. If the true identification error rate in various countries is >0, then current estimates of genetic correlations may be overestimating the true effect of genotype with environment interaction.

The impact of incorrect parentage on international genetic evaluations may also depend on whether the misidentified daughters in a country were sired by only local bulls or by both local and foreign bulls. In the United States, few cows have a foreign sire, except from Canada. Therefore, the errors in this study were introduced mainly into pedigrees of US daughters of US and Canadian bulls but not US daughters of Netherlands or New Zealand bulls. However, the effect of the introduced errors on the correlation estimate between the United States and Canada was the same as the effect on correlation estimates between the United States and New Zealand and between the United States and The Netherlands.

Paternity verification with highly polymorphic DNA markers is technically possible (Ron et al., 1996; Vankan and Faddy, 1999). The practicality of implementation of such verification for the commercial population depends on the costs of extracting DNA and genotyping. Genotyping would indicate whether the sire assigned to a cow is correct. Frequently when a pedigree error is detected, the correct sire may be any bull of a large number of possible candidates. The sire then would be treated as unknown for the purpose of genetic evaluation. When sire replacement is not random, however, and with good farm records on which sires were used in a herd, genotyping may actually assist in identifying the true sire.

In addition to parentage verification costs, the economic returns would be also determined by increased productivity due to genetic progress and increased world trade. Low genetic correlation estimates resulted from misidentified sires would lead to selection bias toward domestic bulls in all countries. Genetic correlations rising towards the biological estimate of genotype by environment interaction would increase exportation of truly outstanding bulls.

CONCLUSIONS

Introducing paternity errors to the pedigrees of 11% of US Holstein cows that were sired by AI bulls had substantial impact on national evaluations and international genetic comparisons. Estimates of inbreeding, genetic trend, sire variance, and across-country genetic correlations decreased. International EBV and bull rankings also were affected. Potential gains from international selection were diminished considerably. When the genotyping of daughters of young sires becomes economically justifiable, large-scale paternity confir-

mation will be possible, and pedigree error rates will decrease.

ACKNOWLEDGMENTS

This research was supported by a USDA grant during the first author's research visit to the Animal Improvement Programs Laboratory, USDA (Beltsville, MD). Permission from USDA, the Canadian Dairy Network (Guelph, ON), the Livestock Improvement Corporation (Hamilton, New Zealand), and NRS (Arnhem, The Netherlands) to use national data is gratefully acknowledged. The authors also thank Drs. Denny Funk, ABS Global, Inc. (DeForest, WI); Kent Weigel, University of Wisconsin (Madison, WI); and Filippo Miglior, Canadian Dairy Network (Guelph, ON) for research review.

REFERENCES

Banos, G., and A. Sigurdsson. 1996. Application of contemporary methods for use of international data in national evaluations. J. Dairy Sci. 79:1117–1125.

Beechinor, J. G., and E. P. Kelly. 1987. Errors of identification amongst cattle presented as progeny of some bulls used in the artificial insemination service in Ireland. Ir. Vet. J. 41:348–353.

Bovenhuis, H., and J. A. M. van Arendonk. 1991. Estimation of milk protein gene frequencies in crossbred cattle with maximum likelihood. J. Dairy Sci. 74:2728–2736.

Christensen, L. G., P. Madsen, and J. Petersen. 1982. The influence of incorrect sire-identification on the estimates of genetic parameters and breeding values. Proc. 2nd World Congr. Genet. Appl. Livest. Prod., Madrid, Spain 7:200–208.

Gelderman, H., U. Pieper, and W. E. Weber. 1986. Effect of misidentification on the estimation of breeding value and heritability in cattle. J. Anim. Sci. 63:1759–1768.

International Bull Evaluation Service. 2000. INTERBULL routine genetic evaluation for dairy production traits, August 2000. http://www-interbull.slu.se/eval/aug00.html. Accessed Oct. 16, 2000.

Israel, C., and J. I. Weller. 2000. Effect of misidentification on genetic gain and estimation of breeding values in dairy cattle populations. J. Dairy Sci. 83:181–187.

Ron, M., Y. Blanc, M. Band, E. Ezra, and J. L. Weller. 1996. Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement. J. Dairy Sci. 79:676–681.

Schaeffer, L. R. 1994. Multiple-country comparison of dairy sires. J. Dairy Sci. 77:2671–2678.

Sigurdsson, A., G. Banos, and J. Philipsson. 1996. Estimation of genetic (co)variance components for international evaluations of dairy bulls. Acta Agric. Scand. 46:129–136.

Vankan, D. M., and M. J. Faddy. 1999. Estimations of the efficacy and reliability of paternity assignments from DNA microsattelite analysis of multiple-sire matings. Anim. Genet. 30:355–361.

Van Vleck, L. D. 1970a. Misidentification and sire evaluation. J. Dairy Sci. 53:1697–1702.

Van Vleck, L. D. 1970b. Misidentification in estimating the paternal sib correlation. J. Dairy Sci. 53:1469–1474.